

Durham Research Online

Deposited in DRO:

18 July 2014

Version of attached file:

Accepted Version

Peer-review status of attached file:

Peer-reviewed

Citation for published item:

Buckle, M. and Chen, J. and Williams, J. (2016) 'Realised higher moments : theory and practice.', *European journal of finance.*, 22 (13). pp. 1272-1291.

Further information on publisher's website:

<http://dx.doi.org/10.1080/1351847X.2014.885456>

Publisher's copyright statement:

This is an Accepted Manuscript of an article published by Taylor Francis Group in *European Journal of Finance* on 19/11/2014, available online at: <http://www.tandfonline.com/10.1080/1351847X.2014.885456>.

Additional information:

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a [link](#) is made to the metadata record in DRO
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the [full DRO policy](#) for further details.

Realised Higher Moments: Theory and Practice

Mike Buckle^a, Jing Chen^a, Julian Williams^{b,*}

^a*School of Management, Swansea University, SA2 8PP, Wales, UK*

^b*Durham University Business School, Durham, DH1 3LB, Scotland, UK.*

Abstract

This paper examines incorporation of higher moments in portfolio selection problems utilizing high frequency data. Our approach combines innovations from the realized volatility literature with a portfolio selection methodology utilising higher moments. We provide an empirical study of the measurement of higher moments from tick by tick data and implement the model for a selection of stocks from the DOW 30 over the time period 2005 to 2011. We demonstrate a novel estimator for moments and co moments in the presence of microstructure noise.

Keywords: Higher Moments, Asset Allocation, Portfolio Management, Co-movement

JEL Classification: G14, G15, G17

1. Introduction and Literature

The importance of higher moments in relation to portfolio selection has been discussed in the literature for some time. For instance, Levy (1969) suggests that expected utility depends on all of the moments of the distribution and that higher moments cannot be neglected. Previous research has also shown that mean-variance portfolio selection techniques can involve a severe welfare loss in the presence of non-quadratic preferences and non-normally distributed asset returns. It is well known that stock returns do not follow a normal distribution. For example, Mandelbrot (1963) and Mandelbrot and Taylor (1967)

*Corresponding Author. Telephone +44 191 33 45301. The authors would like to thank two anonymous referees and the editor Chris Adcock for helpful advice in the production of this paper. Williams acknowledges the support of the European Framework Programme 7 project SECONOMICS.

Email addresses: `m.j.buckle@swansea.ac.uk` (Mike Buckle), `jing.chen@swansea.ac.uk` (Jing Chen), `julian.williams@durham.ac.uk` (Julian Williams)

show that stock returns exhibit excess kurtosis. Fama (1965) finds that large stock returns tend to be followed by stock returns of similar magnitude but in the opposite direction. This can lead to the volatility clustering effect that is related to how information arrives and is received by the market (see Campbell and Hentschel (1992)). This clustering in return volatility has raised a fundamental question on whether a mean and variance asset pricing model using only the first two moments of the return distribution is adequate in capturing variation in average stock returns.

Given that the empirical stock return distribution is observed to be asymmetric and leptokurtic, a natural extension of the two-moment asset pricing model is to incorporate the co-skewness (third moment) and co-kurtosis (fourth moment) factors. An investor whose utility is non-quadratic and is described by non-increasing absolute risk aversion may prefer positive skewness and less kurtosis in the return distribution. Stocks exhibiting negative co-skewness and larger co-kurtosis with the market should therefore be related to higher risk premia. Hence, movement of higher co-moments unfavourable to the investors risk preferences requires compensation in the form of additional returns. A number of empirical studies have shown that investors are willing to accept lower expected return and higher volatility compared to the mean-variance benchmark in exchange for higher skewness and lower kurtosis (see Harvey and Siddique (2000); Dittmar (2002) and Mitton and Vorkink (2007)).

Recent contributions to the literature by de Athayde and Flores (2004), Jondeau and Rockinger (2003), Harvey et al. (2010) and Cvitanic et al. (2008) have looked at the inclusion of the first four co-moments into the distribution of portfolio returns and how to build 3-4 moments frontiers and select appropriate portfolios from these frontiers. There have been some other attempts to model portfolio selection taking into account third (see Mencía and Sentana (2009)) and third/fourth moments (see Jondeau and Rockinger (2006) and Martellini and Ziemann (2010)). The main problem with these models is that they generally use weekly or monthly returns to estimate the parameters. For estimation of coskewness or cokurtosis parameters, there is a severe dimensionality problem. For example, optimizing a portfolio of 20 stocks would require estimation of 210 variance-covariance parameters,

1,540 skewness-coskewness parameters and 8,855 kurtosis-cokurtosis parameters. This makes efficient implementation of portfolio selection unrealistic. With low frequency data there is likely to be insufficient data available. One solution to this is to use techniques that reduce the number of parameters to be estimated. In this paper we adopt an alternative approach. This approach draws upon the literature relating to high frequency construction of realized volatility and covariance. Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002) have shown that the sum of squared returns converges to the unobserved integrated volatility as the intraday-interval goes to zero. A weakness with this approach to modelling volatility is that it can be very sensitive to market frictions when applied to returns recorded over very short time intervals such as 1 minute or less. Barndorff-Nielsen et al. (2008a,b) propose realized kernel estimators of quadratic variation that are robust to certain types of frictions.

Following the highly volatile market conditions observed in many asset markets during the 2007-2011 period, more reliable methods of computing the ex-post variation in asset returns have been called for. The importance of correct measurement of moments and comoments and correctly dealing with contaminants in empirical data cannot be overstated. The Basel III framework has required more robust methods of computing distributions of asset prices in response to high levels of ‘tail-risk’ observed in many bank portfolio holdings. Utilizing the higher moments and co-moments has been suggested as a means of overcoming the perceived non-Gaussian properties of financial asset returns.

With automated trading accounting for nearly 70% of trades on the NYSE (see O’Hara (2003), providing approaches for computing multivariate density functions is becoming ever more important). Using sequences of moments and comoments is generally as quick or quicker to execute than most other semi or non-parametric methods. Moments and comoments are generated by the r^{th} order sequence of polynomial expansions of a density function, where $r \in \mathbb{N}^+$. In the case of multivariate distributions the moments are formed by the r^{th} order expansion of the vector function into a space of dimension r . Recent examples of the synthesis of measured moments and decision making can be seen in Mencía and Sentana (2009) for the building of skew portfolios by mixtures of normals; and in Briec and Kerstens (2010), who address the construction of mean-variance-skewness portfolios

(moments 1, 2 and 3) using a shortage function approach.

The mean-variance framework that is at the heart of most financial models such as the Sharpe-Lintner CAPM, is a case of utilizing the first and second moments. However, expansion to a higher moment framework is non-trivial: the normal/multi-normal distribution has the advantage of having the first and second moments as the parameters that characterize the univariate/multivariate density probability function. After expansion beyond the second moment no other univariate or multivariate functions have this property and as such parametrization and confidence testing using the normal sample rules, convergence limits and likelihoods become very difficult.

From the perspective of portfolio selection the problem of the inclusion of higher moments also has many difficulties. Brockett and Garven (1998) demonstrate that arbitrary truncation of the polynomial expansion of a utility function to match measured portfolio moments may produce sub-parsimonious decision outcomes, i.e. an investor when faced with two portfolios may find that expected utility maximization may occur by choosing a portfolio with lower mean, higher variance and lower skewness over an alternative, obviously negating the standard framework of preferences for higher return, lower variance and higher positive skewness. This is due to the structure of the truncation; once the decision has been made to include higher moments, the truncations of the polynomial expansions may no longer have neat zero remainders and as such domains of decisions exist whereby decision making deviates from those that would seem obviously appropriate. Key to this issue is the structure of the hyper-surface that defines the efficient frontier (the constrained opportunity set) and the hyper-surface defined by the available preferences. Time variation in the higher moment structure will require substantial adjustment to the equilibrium allocation.

This paper introduces tractable approaches to estimating and incorporating higher moments in standard financial management settings using ultra-high frequency data. More importantly, we bring together the literature on higher moments in portfolio selection and high frequency realized variance estimation techniques. This provides a new approach to optimal portfolio selection that incorporates information in higher moments and at the same time allows for efficient estimation of the moment parameters. Our results demonstrate the

inclusion of directly estimated higher moments radically improves the fit of forecasted return densities.

The remainder of the paper is organised as follows: §(2) explains our notation approach and outlines the general theory of measurement for moments and comoments of asset returns; §(3) introduces the idea of measuring comoment arrays for ultra-high frequency data and reports a monte-carlo case study of our comoment estimator. We also present a case study on measuring the distribution of a simulated portfolio of 20 selected stocks from the Dow 30 around the Lehmen Brothers Chapter 11 event in September 2008. §(4) presents our main set of empirical results testing the moment-comoment sequence for a selection of 20 stocks from the Dow 30. We present some brief conclusions and ideas for further research in §(5).

2. The Co-Moment Structure of Asset Returns

We start with a vector of log prices denoted $\mathbf{y}(t)$ updated at irregular frequencies index by the vector process \mathbf{t} , setting $x_{i,t} = y_{i,t+1} - y_{i,t}$. This process exhibits realised co-products for moment two: $\langle x_i, x_j \rangle_T = \int_0^T m_{i,j} dt$, three $\langle x_i, x_j, x_k \rangle_T = \int_0^T m_{i,j,k} dt$, four $\langle x_i, x_j, x_k, x_l \rangle_T = \int_0^T m_{i,j,k,l} dt$ and so on until the r^{th} moment $\langle x_{[r]} \rangle_T = \int_0^T m_{[r]} dt$, where $[r]$ denotes general covariant index of length r of the contravariant vector \mathbf{y} .

Each measured comoment is then stored in an r dimensional array with elements defined by $\langle x_{[r]} \rangle_T = \sum_{t_i=1}^{T_i} (y_{i,t+1} - y_{i,t})^{[r]} \equiv \mathbf{x}^{[r]}$. Constructing this array presents two problems: first is indexation; multidimensional arrays do not inherently lend themselves to portfolio allocation and risk management problems. Second is the time index; for n stocks each stock is updated at asynchronous points as such the time indexation is problematic. We shall deal with each of these points in turn in the remainder of 2 and 3.

Measuring Comoments

Let us denote the array of cumulative expected comoment for a time interval $0, T$ as $\mathbb{E}(\mathbf{x}^{[r]}) = \mathcal{M}_r$, where \mathcal{M}_r is an r array of co-moments and we specify $\mathbf{x}^{[r]}$ as the vector permuted outer product of the vector \mathbf{x} with itself; e.g. for the second co-moments, this changes to the transpose operator designated T i.e. $\mathbf{x}^{[2]} \equiv \mathbf{x}\mathbf{x}^T$. We shall also make a slight

abuse of notation by using the expectation operator to represent the ex-post realisation of the comoments, therefore $\mathbb{E}(\mathbf{x}^{[r]}) = \sum_{t=1}^T \mathbf{x}_t^{[r]}$.

All moments are assumed to be un-normalised and un-centered. The resultant array \mathcal{M}_r is an r dimensional array (classical tensor), with r^{th} order super-symmetry. In matrix notation we can define a flat matrix, containing the identical elements to \mathcal{M}_r , using the higher *vec* operator we can define the relation: $vec(\mathbf{M}_r) = vec(\mathcal{M}_r)$. In this case the *vec* operator stacks the fibre bundles of the array \mathcal{M}_r , in such a way that the inverses are homeomorphic, i.e. for a pair of inverse transformations $ivec_r, ivec_2$ of an array, the following definitions hold,

$$ivec_r(vec(\mathbf{M}_r)) = \mathcal{M}_r \quad ivec_2(vec(\mathcal{M}_r)) = \mathbf{M}_r \quad (1)$$

We now define the matrix \mathbf{M}_r in terms of Kronecker powers and first order permutations, (transposes) of \mathbf{x}

$$\mathbb{E}((\mathbf{x}^{[\otimes r-2]}) \mathbf{x}^T) = \mathbf{M}_r \quad (2)$$

This allows the rewriting in Kronecker power notation of the comoment array as $\mathbf{x} = \mathbf{x}^{[\otimes 0]}$. For a given vector of weights, ω , then the weights array is defined as $\omega^{[r]} = \mathcal{W}_r$, and we can specify the following *matricizing* condition, $vec(\mathbf{W}_r) = vec(\mathcal{W}_r)$, or, $(\omega^{[\otimes r-2]}) \omega^T = \mathbf{W}_r$. From this definition r^{th} moment μ_r of a random variable $x = \omega^T \mathbf{x}$, is therefore $\mu_r = \mathcal{W}_r \bullet \mathcal{M}_r$. Where \bullet is the inner tensor product of two identical arrays. In vector notation it can be written as,

$$\mu_r = vec(\mathbf{W}_r)^T vec(\mathbf{M}_r) \equiv vec(\mathcal{W}_r)^T vec(\mathcal{M}_r) \quad (3)$$

Subsequently the r^{th} moment of x is then defined in matrix notation as,

$$\mu_r = vec((\omega^{[\otimes r-2]}) \omega^T)^T vec(\mathbb{E}((\mathbf{x}^{[\otimes r-2]}) \mathbf{x}^T)) \quad (4)$$

The first 4 co-moments of \mathbf{x}^1 .

Moment Array by Vector Permutations	Moment Matrix, in Kronecker Product Notation	Moment Matrix, in Kronecker Power Notation
$\mathbb{E}(\mathbf{x}^{[1]}) = \mathcal{M}_1$	$\mathbb{E}(\mathbf{x}) = \mathbf{M}_1$	$\mathbb{E}(\mathbf{x}) = \mathbf{M}_1$
$\mathbb{E}(\mathbf{x}^{[2]}) = \mathcal{M}_2$	$\mathbb{E}(\mathbf{x}\mathbf{x}^T) = \mathbf{M}_2$	$\mathbb{E}((\mathbf{x}^{[\otimes 0]}) \mathbf{x}^T) = \mathbf{M}_2$
$\mathbb{E}(\mathbf{x}^{[3]}) = \mathcal{M}_3$	$\mathbb{E}((\mathbf{x} \otimes \mathbf{x}) \mathbf{x}^T) = \mathbf{M}_3$	$\mathbb{E}((\mathbf{x}^{[\otimes 1]}) \mathbf{x}^T) = \mathbf{M}_3$
$\mathbb{E}(\mathbf{x}^{[4]}) = \mathcal{M}_4$	$\mathbb{E}((\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}) \mathbf{x}^T) = \mathbf{M}_4$	$\mathbb{E}((\mathbf{x}^{[\otimes 2]}) \mathbf{x}^T) = \mathbf{M}_4$
	\Downarrow	
$\mathbb{E}(\mathbf{x}^{[r]}) = \mathcal{M}_r$	$\mathbb{E}((\mathbf{x} \otimes \mathbf{x} \otimes \dots \otimes \mathbf{x}) \mathbf{x}^T) = \mathbf{M}_r$	$\mathbb{E}((\mathbf{x}^{[\otimes (r-1)]}) \mathbf{x}^T) = \mathbf{M}_r$

Weighting the empirically observed co-moments yields

$$\begin{aligned}
\mu_0 &= 1 \\
\mu_1 &= \omega^T \mathbf{x} \\
\mu_2 &= \text{vec}(\omega \omega^T)^T \text{vec}(\mathbb{E}(\mathbf{x}\mathbf{x}^T)) \equiv \omega^T (\mathbb{E}(\mathbf{x}\mathbf{x}^T)) \omega \\
\mu_3 &= \text{vec}((\omega \otimes \omega) \omega^T)^T \text{vec}(\mathbb{E}((\mathbf{x} \otimes \mathbf{x}) \mathbf{x}^T)) \\
\mu_4 &= \text{vec}((\omega \otimes \omega \otimes \omega) \omega^T)^T \text{vec}(\mathbb{E}((\mathbf{x} \otimes \mathbf{x} \otimes \mathbf{x}) \mathbf{x}^T)) \\
&\Downarrow \\
\mu_r &= \text{vec}((\omega^{\otimes (r-1)}) \omega^T)^T \text{vec}(\mathbb{E}((\mathbf{x}^{\otimes (r-1)}) \mathbf{x}^T))
\end{aligned} \tag{5}$$

We can compute correlation analogues, denoted for moments 3 and 4 as skew-correlation and kurtic-correlation arrays/matrices using the same approach as we do for quadratic correlation matrices. For an array \mathcal{M}_r let \mathbf{m}_r denote the r root of the n vector of of the super diagonal, i.e. for a co-skewness array the elements of \mathbf{m}_3 are $(\langle x_i, x_i, x_i \rangle_T)^{\frac{1}{3}} = \sqrt[3]{\sum_{t_i=1}^{T_i} (x_{i,t+1} - x_{i,t})^3}$.

¹We utilize T, to represent the simple conjugate transpose operator, vec represents the column-wise stacking of a matrix or array, to a column vector, we set $\mathbf{a}^T \mathbf{b} = \sum_{i=1}^N a_i b_i$, for two N length vectors \mathbf{a} and \mathbf{b} .

For an array $\mathcal{H}_r = \mathbf{m}_r^{[r]}$, the high comoment correlation array is derived from the relation: $\mathcal{M}_r = \mathcal{H}_r \times \mathcal{R}_r$, where \times denotes element by element multiplication. In matrix notation this simplifies to:

$$\mathbf{M}_r \div \mathbf{H}_r = \mathbf{R}_r \quad (6)$$

where \div is the element by element division operator of two identically sized arrays. We can finally expand this to a Kronecker power notation to yield:

$$\mathbf{M}_r \div (\mathbf{m}_r^{[\otimes(r-1)]} \mathbf{m}_r^T) = \mathbf{R}_r \quad (7)$$

This is a useful identity as we can use the array \mathbf{R}_r to isolate only the ‘off super diagonal’ changes in skew-correlation and kurtic-correlation.

Following Guidolin and Timmermann (2008) we can write the objective function of a portfolio optimisation problem in terms of an expanded utility function. Let $u(c)$ be a r differentiable utility function, fixing a particular investment point c and setting $u^{(j)}$ to be the j derivative of u at c , the asset allocation problem is

$$\max_{\omega} \sum_{j=1}^r (j!)^{-1} u^{(j)} \sum_{t_i=1}^{T_i} \omega^{[j]} (x_{i,t+1} - x_{i,t})^{[j]} \quad (8)$$

under our notation this is equivalent to:

$$\max_{\omega} \sum_{j=1}^r (j!)^{-1} u^{(j)} \text{vec}((\omega^{\otimes(j-1)}) \omega^T)^T \text{vec}(\mathbb{E}((\mathbf{x}^{\otimes(j-1)}) \mathbf{x}^T)) \quad (9)$$

several recent contributions have demonstrated for low frequency data that the inclusion of higher moments in the portfolio optimisation substantially changes the actual asset allocation, see for instance Harvey et al. (2010). In addition to expanding the asset allocation problem to higher dimensions the moments of the portfolio can be used to construct the density function characterising the portfolio.

A simple method of reconstructing the portfolio density function from its moments is the von Mises (1947) stepwise approach, which proposes matching the moment sequence

to the moments of some arbitrary density function. We can mix a variety of distributions together to exactly match the empirical distribution sequence, for instance the half/log normal, Weibull, Gamma, Maxwell, Exponential, Chi-Squared, Rayleigh, Hypergeometric, Cauchy, Student and F -distributions offer an eclectic library. In addition to using alternative distributions we can also mix distributions of the same type with differing parameters. The first step is to deal with the information content that the moment sequence contains, for a Hankel matrix Δ_r , partitioned into sequences of moments as follows,

$$\Delta_{2r} = [\mu_{[0,r]}, \mu_{[1,r+1]}, \dots, \mu_{[r,2r]}]^T \quad (10)$$

A new partitioned matrix, Φ may now be formed,

$$\Phi = \begin{bmatrix} \mu_{[0,r]} & \mu_{[0,r]} & \cdots & \mu_{[0,r]} & \mu_{[0,r]} \\ \mathbf{0} & \mu_{[1,r+1]} & \cdots & \mu_{[1,r+1]} & \mu_{[1,r+1]} \\ \mathbf{0} & \mathbf{0} & \ddots & \vdots & \vdots \\ \vdots & \vdots & \cdots & \mu_{[r,2r-1]} & \mu_{[r,2r-1]} \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mu_{[r,2r]} \end{bmatrix}^T \quad (11)$$

where,

$$\mu_{[p,q]} = [\mu_p, \dots, \mu_q]^T \quad (12)$$

For a vector of coefficients $\mathbf{c} = [c_0, c_1, c_2, \dots, c_{2r}]^T$ the following linear algebra problem is then solved as a method of moments problem,

$$\Phi \mathbf{c} + \mu_{[r+1,2r+1]} = \mathbf{0} \quad (13)$$

To obtain the abscissa values, $\{\nu_1, \nu_2, \dots, \nu_{2r-1}\}$, the roots of the $2r - 1^{th}$ polynomial with coefficients described by \mathbf{c} are evaluated. This system maybe rewritten as a set of Legendre polynomials, see Devroye and Hwang (2006). We can then compute approximations of the critical boundaries allowing for the computation of value at risk and expected shortfall

measures. This is therefore analogous to a method of moments problem if we set the target moments as a mixture of the distributions described previously.

The target moments can be derived from a single characteristic function or set of characteristic functions as required. One consequence of this is that if the multivariate moments of the individual assets is very complex then the portfolio planner can in effect choose the class of distribution of the portfolio rather than simply adjust parameters within a single family of distributions.

3. Time Distortion Effects

The second issue in dealing with ultra high frequency data is the effect of randomised time intervals for stock updates creating time distortion effects. The ‘Epps’ effect introduced in Epps (1979) illustrates that correlations are suppressed as sampling intervals increase. Recent work by Aït-Sahalia et al. (2011) has indicated that microstructure noise frictions also play a part in either dampening or exacerbating the level of volatility. Our approach uses a multi timescale framework in line with the ‘needlework’ estimator of Audrino and Corsi (2010) and Barndorff-Nielsen et al. (2009).

The issue with estimating cross variational products with asynchronous updating is that a very fine (say second by second) regularised grid such as that used in Williams et al. (2012) often results in many zero price changes. This can have two effects, first it can make covariance matrices that are near singular and hence not invertible thus useless for portfolio management and second suppress absolute levels of correlation, by adding large numbers of zero values to sums of products.

For instance there are 78 five minute blocks of time in an NYSE trading day, therefore there are 23,400 seconds in a grid, which is one order of magnitude more than the average number of transacted price updates for most stocks in our sample².

Consider an n vector of stocks with tick intervals $t_i \in \mathbf{t}$, the objective is to compute the cross product $\langle x_{[r]} \rangle_T = \sum_{t_i=1}^{T_i} (x_{i,t+1} - x_{i,t})^{[r]}$, for some time interval t, T , which we shall

²In this study we look at transacted prices, including bid-ask quotes and re-quotes does not necessarily mitigate this issue

assume to be one day. Each stock has its own clock and update times labelled t_1, \dots, t_T , where t_T is assumed to be the last update prior to the end of the day. To estimate the cross products a set of tick times needs to be built such that the maximum amount of actual cross variation is captured.

The ‘needlework’ estimator operates by assigning changes in prices in one stock to the closest prior price change in another stock. Therefore the ‘master’ clock of tick-times are not uniform in their sampling frequency. We take this a step further by adding the multi timescale concept of Aït-Sahalia et al. (2011) to improve our sampling performance.

Consider a set of fixed regular time intervals indexed by \tilde{t} , for instance 5 minute intervals. For a cross section of stocks there will be a set of updated prices (ticks), $\{\tilde{t}_j < t_{i \in \{1, n\}} < \tilde{t}_{j+1}\}_{\tilde{t}_j}^{\tilde{t}_{j+1}}$. Assuming that the update frequency is some form of point process then there will be a distribution of updated prices over the interval. If we use the most frequently updated stock then many other stocks will report zero price changes relative to this clock.

Setting $N_i(\tilde{t}_j, \tilde{t}_{j+1})$ as the counter for the number of updated prices for the i stock, the chosen clock for a cross section of n stocks is the set of tick times $\{\tilde{t}_j < t_{i^*} < \tilde{t}_{j+1}\}_{\tilde{t}_j}^{\tilde{t}_{j+1}}$, whereby

$$i^* = \arg \min_{i \in \{1, n\}} N_i(\tilde{t}_j, \tilde{t}_{j+1}) \quad (14)$$

i.e. the stock with lowest number of updated ticks over the interval. Our adaptation of the needlework estimators works well for cross sections of stocks with frequent updates such as those traded on exchanges without substantial execution lags. For instance the NASDAQ permits 30 second delays before reporting of executed trades (see Aït-Sahalia et al. (2011)) therefore the mixing of delayed trades with instantly executed trades would create some distortions at our ultra high frequency sampling range.

Monte-Carlo Study Moments 2, 3 and 4

To test our cross product estimator we construct a Monte-Carlo study. We take a twelve stock portfolio, in the authors experience this is a typical size of cross section for a single investment book for a statistical arbitrage hedge fund. Returns are assumed to be generated by a copula based distribution with skewed marginals to create higher co-moments.

The distributional assumptions are as follows. Let the intraday underlying log price process be a vector random walk $\delta_T = \delta_0 + \sum_{t=1}^T \delta_0$ where the revealed price increment is $\delta_t \sim \mathfrak{F}(\theta)$. Where $\mathfrak{F}(\theta)$ is a multivariate density function, with parameters θ . Using Sklars theorem we decompose $\mathfrak{F}(\theta)$ into a joint distribution coupling distribution, $\mathfrak{T}(\theta^{copula})$ and independent marginal distributions, $\mathfrak{F}^i(\theta^i)$. In our case we shall use a multivariate T distribution, with parameters $\theta^{copula} = \{\nu, \Omega\}$ to impart second and fourth co-moments and a generalised extreme value distribution to create skewness in the marginals, $\theta^i = \{k^i, \mu^i, \sigma^i\}$. The motivation for this choice comes from the literature on non-Gaussian approaches to approximating asset returns, see for instance Rachev and Mitnik (2000) for an overview.

The ‘master’ clock for the underlying process is assumed to be a uniform grid, t refreshed at a rate such that $\Delta t < \Delta t_i, \forall i \in n$, i.e. faster than the fastest refresh rate for the individual tick times. In our case this is 234,000 gridded price changes a day, equivalent to a one tenth of a second uniform grid. Figure 1 presents an illustration of our estimator sampling approach versus a fixed grid.

Individual ticks t_i are refreshed via a poisson point process with rate parameter λ^i . We assume that λ^i is distributed across stocks, but is constant for an individual stock for the t, T interval. Previous research such as Admati and Pfleiderer (1989) have suggested that more trades will take place in early trading, however we assume that our analysis will exclude the opening call auction period.

Therefore the Monte-Carlo algorithm works as follows, construct a vector sequence of draws of δ_t at the master clock refresh rate. Each draw is constructed by drawing a vector from a multivariate T distribution with chosen degrees of freedom ν and positive definite parameter matrix Ω . Each marginal draw is then transformed by a cumulative univariate T distribution function, with the same degrees of freedom ν and then transformed again by the inverse cumulative generalised extreme value distribution with chosen shape parameter k^i and centering and spread parameters μ^i and σ^i , resulting in a marginal draw of δ_i . The master clock draw of δ_t has its empirical moments and comoments computed and these are set as the objective moments for the sampling procedure.

Next, we construct a sequence of ultra high frequency refresh times for the individual

stocks. We draw a master set of λ^i , from a normal distribution, such that the average is 5,000 or 8,000 ticks per day and the standard deviation is 250, equivalent to the observed spread of refresh times for prices on the NYSE on a high versus low activity trading day. A set of refresh times are drawn for the day and the prices are sampled from the nearest preceding time increment for the vector draw from δ_t , which we designate $\tilde{\delta}_{t_i}^i$. We run a set of experiments ranging from where the observed log prices $x_{t_i}^i = \tilde{\delta}_{t_i}^i$ have no extra contaminants to a point where we add an increasing amount of i.i.d. microstructure noise $x_{t_i}^i = \tilde{\delta}_{t_i}^i + \epsilon_{t_i}^i$, $\epsilon_{t_i}^i \sim \mathcal{N}(0, \gamma_i^2)$.

We then compare three coproduct estimators for the second, third and fourth moments and comoments. First a uniform grid of ten seconds, next the simple ‘needlework’ estimator of Audrino and Corsi (2010) and finally our ‘adjusted needlework’ estimator. For a given moment-comoment array let \mathbf{M}_r be the ‘true’ array and let $\tilde{\mathbf{M}}_r$ be a candidate matrix computed via one of the the methods mentioned previously, then the estimated error is computed by $\tilde{\Lambda}_r = \left\| \mathbf{M}_r - \tilde{\mathbf{M}}_r \right\|$, where $\|\cdot\|$ is the Frobenius norm of a matrix.

Table 1 presents the results of the Monte-Carlo study. The columns of Table 1 define the type of estimator used whilst the rows present the various different input configurations of the simulation. The proportions of γ are relative to the level of variance of the underlying price process, therefore 0.01, means that the microstructure noise is equivalent to 1% of the quadratic variation in the actual price process.

It is apparent that there are trade-offs in the selection of the various approaches, based on the assumed magnitude of the variance of the unobserved microstructure noise is. Speed should also be commented on here, as the adjusted needlework estimator uses existing tick times rather than searching for updated ticks across the cross section it is approximately one order of magnitude faster, albeit not as fast as the fixed grid estimators. Another point to notice is that the higher moments are always more accurately predicted than the covariance and suffer far less deterioration from increasing microstructure noise.

[INSERT TABLE 1 HERE]

[INSERT FIGURE 1 HERE]

Example: 20 Stocks From the Dow Jones During August, September and October 2008

We now move on to illustrate the effect of including higher comoments in the evaluation of portfolio density functions. We will compute the density function for a portfolio of twenty stocks using either the first two or the first four moments for a log utility maximising investor. For our example we collect every recorded transaction on the NYSE for the 20 stocks from eight weeks prior to the default of Lehman Brothers to eight weeks after this event. We choose 20 out of the 30 stocks from the Dow Jones, as a 20 variate model is computationally tractable for a long time series. As of 2011, the companies in the sample constitute roughly one quarter of the total market capitalisation for all US equities.

[INSERT TABLE 2 HERE]

Table 2 presents the list of stocks, their RIC codes and the number of updated (transacted) ticks for the period of January 1, 2005 to October 10, 2011. The stocks are collected into a portfolio that maximises log-utility using the approach suggested in formula 9. The expected utility is maximised numerically using a standard sequential quadratic optimisation algorithm. The weights are set using the data from the four and eight weeks prior to September 16, 2008 and carried over into the four and eight weeks afterwards. We use two different windows to provide a comparison (See Figure 2).

We utilise a library of distribution functions to minimise the objective moments. For the two moment system we use a Gumbell distribution with analytic moments as the target distribution for the portfolio. When including the next four moments more general distributions are available, we choose a generalised extreme value distribution from the Fréchet family as the target distribution. Again the target moments are analytic and easily fitted to those measured from the high frequency data.

[INSERT FIGURE 2 HERE]

Figure 2 presents our two experiments using the four or eight weeks prior to predict the density for the four or eight weeks after. The striking appearance is the substantial variation in portfolio densities from either choice of moment or time period.

It is quite obvious to see that the chapter 11 announcement of Lehmen Brothers on September 16, 2008 has a marked effect on the elements of the twenty stock portfolio. There are dramatic increases in the levels of variance, and there is also a noticeable decrease in the level of skewness and kurtosis observed.

For the portfolio density functions we can see that the inclusion of moments 3 and 4 in the decision system markedly improve the level of fit compared to the estimates of moments 1 and 2 post September 16, 2008. The higher moments capture a far higher level of downside risk in the four weeks after September 16, 2008 and correctly predict that the portfolio of twenty stocks could have a 50% chance of reducing in value by 50% over the next four weeks.

The useful aspect of this exercise is in illustrating the benefits of including higher moments in outlining the complex adjustments in the distribution of returns during a time of market stress. The mean variance approach, cannot capture the extreme values that the portfolio can take during this time of market stress. In fact we start to see agglomeration of probability mass at catastrophic levels near the origin (i.e. the portfolio having zero value).

4. Empirical Application and Data

We run the rolling comparative forecast test for the second, third and fourth realised higher moments for the 20 stocks utilised in the portfolio example in §§(3), for the 1,706 days from January 1, 2005 to October 10, 2011. A summary of the total number of observations is given in the third column of Table 2. The data is partitioned by day and each day is processed using the adjusted ‘needlework’ approach into a standardised set of time stamps. The complete standardised dataset is available from the authors website.

For each day returns are computed using the approach outlined in §§(3) and the second, third and fourth moments and comoments are then computed for each day. For the long run forecasts we use a 66 day window, which is the number of trading days in a quarter. The

total number of processed, ‘needlework’, updated ticks is 7,259,030 or an average of 4,255 updated transactions a day.

For one stock, TRV Travelers Companies, for four days (January 3-6, 2006) from the sample there are no trades recorded. We replace the missing data with transactions from the previous week to maintain a consistent time series. For the comparative analysis the loss function is set as the first g predictor to be the previous days moment and comoment sequence and the second h predictor to be a running quarterly average of 66 days.

Discussion with practitioners suggested that a one quarter averaging risk assessment versus a day on day trading book was a good comparison for this exercise.

Results and Analysis

Table 2 presents the descriptive statistics for the numbers of ticks available for each day for the 20 stocks in the sample for each year in the sample from 2005 to 2011, recalling that the sample finishes in October, 2011. The total number of ticks for the each of the stocks in question varies between 10.7 million for United Technologies and 22.3 million for General Electric. For the Log utility portfolio used in the example seven stocks are short and thirteen are long, with the largest weight being placed in Johnson and Johnson.

The stocks are amongst the most actively traded in the whole world and are generally as actively traded or more actively traded than tracker paper such as the S&P depository receipt. This makes them good candidates for this type of study.

Table 3 presents breakdown by year of daily descriptive statistics for the number of ticks for cross section of the sample. The table illustrates the impact of the ‘sub-prime’ and ‘credit crunch’ financial crises on average trading for 2007-2009 with average numbers of daily ticks increasing from 5,552 in 2006 to 12,717 in 2008.

The table also reports the average, median, minimum and maximum number of zero returns recorded daily for each year for the cross section. Recall that the total number of tick changes is the mean number of ticks times twenty, therefore for 2009, the average for the daily total number of ticks is 210,286 and there was on average 22,941 zero returns or less than 10% of returns were zero. The maximum number of zero returns recorded was 48,318

for a single day in 2007. This reinforces the practitioner view that the Dow 30 components are amongst the most actively traded and continuously updated assets available in the world and an excellent case study for our estimator.

[INSERT TABLE 3 HERE]

For the measured higher moments we have divided the results into two parts, first the $j \in 1, \dots, r$ -root of the diagonal elements which we designate as $\mathbf{m}_{1,\dots,r}$. For each moment array there are n super diagonal elements. For each year we have tabulated the cumulative value of $\mathbf{m}_{1,\dots,r}$, for the second, third and fourth moments (variance, skewness and kurtosis) and averaged over the cross section of stocks. These results are reported in Table 4.

The obvious impact of the financial crisis is on the annualised volatility of the stocks in the sample, during 2008 we see the highest second moment increase from 25% to over 75%. The average level of Skewness is more stable over the sample period however the maximum realised skewness increases in 2008 by three fold from around +2 to +8. The lowest level of skewness is almost perfectly symmetrical to the highest level of observed skewness.

The maximum level of raw kurtosis increases to +19 in 2008 from 4.5 in 2005 three years earlier. This level of kurtosis is massively in excess of the level to be expected from the corresponding level of variance, i.e an excess of $(18.99^4 \times 0.77^{-4} \gg 1)$ over a normal distribution.

[INSERT TABLE 4 HERE]

Table 5 presents the descriptive statistics for the cumulative annualised and normalised off-diagonal elements for the moment arrays for the covariance-correlation, coskewness-skew-correlation and cokurtosis-kurtic-correlation. For the second moment correlations 2008 had the highest minimum level of correlation at 21% and the second highest maximum correlation at 62%. More interestingly the maximum level of skew-correlation peaks at 1.4 (note that the normalised coskewness is not constrained between -1 and 1). This is a very high value and

indicates that certain stocks as three-tuples exhibit very high levels of positive codependence. The small size of the maximum and mean kurtic-correlation in 2008 is actually somewhat smaller than in 2009-2011. The co-risk and co-dependence have substantially increased after 2008 and this tail dependence has not reverted back after the market stress of 2008.

Figures 3 to 5 present the plot of the average of the $j - root$, for $j \in 1, \dots, r$ of the daily cumulative diagonal moments. The dotted line presents the actual measured daily moments and the unbroken black line is a ten day moving average.

We see that after 2008 that the daily ex-post level of $n^{-1} \sum \mathbf{m}_2$ rises substantially. What is more interesting is the effect on the level of skewness. On average pre 2008 the daily level of skewness is relative low, however after 2008 we see sudden spikes and troughs in the level of skewness as the market shifts suddenly.

The level of kurtosis also spikes in 2008, however the overall level of kurtosis is always far in excess of that predicted under a normal distribution. However the level of excess kurtosis during 2008 is very large and this is also visible in the annual aggregates in Table 5.

[INSERT FIGURES 3 TO 5 AS A SINGLE PANEL HERE]

Figures 6 to 8 present the plot of the average of then normalised daily off diagonal elements for the covariance $offdiag\mathbf{R}_2$, coskewness $offdiag\mathbf{R}_3$ and cokurtosis $offdiag\mathbf{R}_4$. Again we see substantial differences in the time pattern of the daily aggregates for instance the covariance and cokurtosis exhibit noticeable rises after 2008 but the variation in skewcorrelations is far lower. This is a new result in the literature and suggests that for this case study sample the crisis impacts returns in a symmetrical manner, via variance-covariance and kurtosis-cokurtosis. Skewness only plays a part in the most extreme circumstances, during a short period around 2008. The policy implications of this observation are quite interesting as it suggests that the market can continue to function in most cases (asymmetric returns indicate substantial increases in information asymmetry, see for instance the theoretical work of Admati and Pfleiderer (1988)). The high levels of coskewness seen during the crisis indicates

that during the crisis the market ceased to function in anything resembling the ‘orderly market’ that the securities and exchange commission (SEC) states as being its regulatory target³.

The usefulness of the inclusion of realised higher moments is in fully populating a density function that does not coverge as quickly as a standard Brownian motion. Once the diffusion underlying asset prices deviates from this convention then sampling at lower frequencies will miss many of the important underlying properties of the data generating process.

[INSERT FIGURES 6 TO 8 AS A SINGLE PANEL HERE]

[INSERT TABLE 5 HERE]

5. Concluding remarks

This paper has introduced a method to utilize higher moments empirically estimated from high frequency data in the asset allocation problem. The proposed method utilizes the empirically estimated co-moments and matches them to those from a generalized density function that may be composed from a mixture of distributions. Within the expected utility maximizing framework it is demonstrated that combining the higher moments into a coherent framework improves portfolio performance as information about higher order moments makes an important contribution to portfolio choice. The empirical results show an increase in variance and kurtosis at the time of the 2007/2008 crisis. Although there is some spiking in skewness in 2008 when looking at the smoothed series of the third moment, there is very little increase in 2008 when compared to the second or fourth moment. During the 2007/2008 financial crisis, covariance and cokurtuosis increase when the crisis occurs and remain high for a prolonged period afterwards. In contrast, coskewness only increases at the most extreme point of the crisis in 2008, and then reverts back to the pre-crisis levels. This approach appears to be a relatively simple method for increasing the information set

³See SEC documentation: <http://www.sec.gov/about/whatwedo.shtml>

used in asset allocation problems and for dealing explicitly with observed deviations from multivariate-normality and the associated moment sequence that appear in many asset return series. The methodology can be applied to time-varying conditional moments, albeit at a far greater computing cost, despite restricting the number of dimensions to the elements of the super-diagonal.

References

- Admati, A. R. and P. Pfleiderer (1988). A theory of intraday patterns: Volume and price variability. *The Review of Financial Studies* 1(1), 3–40.
- Admati, A. R. and P. Pfleiderer (1989). A theory of intraday and day-of-the-week mean effects. *The Review of Financial Studies* 2(2), 189–223.
- Aït-Sahalia, Y., P. A. Myklandb, and L. Zhang (2011). Ultra high frequency volatility estimation with dependent microstructure noise. *Journal of Econometrics* 160(1), 160–175.
- Andersen, T. G., T. Bollerslev, F. X. Diebold, and H. Ebens (2001). The distribution of realized stock return volatility. *Journal of Financial Economics* 61(1), 43–76.
- Audrino, F. and F. Corsi (2010). Modeling tick-by-tick realized correlations. *Computational Statistics & Data Analysis* 54(11), 2372–2382.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008a). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica* 76(6), 1481–1536.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008b). Multivariate realised kernels: consistent positive semin-definite estimators of the covariation of equity prices with noise and non-synchronous trading. Technical Report 397, Economics Series Working Papers, University of Oxford.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realized kernels in practice: Trades and quotes. *Econometrics Journal* 12(3), C1–C32.
- Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 64(2), 253–280.
- Briec, W. and K. Kerstens (2010). Portfolio selection in multidimensional general and partial moment space. *Journal of Economic Dynamics and Control* 34(4), 636–656.
- Brockett, P. L. and J. R. Garven (1998). A reexamination of the relationship between preferences and moment orderings by rational risk averse investors. *The Geneva Papers on Risk and Insurance Theory* 23(2), 127–137.
- Campbell, J. Y. and L. Hentschel (1992). No news is good news: an asymmetric model of changing volatility in stock returns. *Journal of Financial Economics* 31(3), 281–318.
- Cvitanic, J., V. Polimenis, and F. Zapatero (2008). Optimal portfolio allocation with higher moments. *Annals of Finance* 4(1), 1–28.
- de Athayde, G. M. and R. G. J. Flores (2004). Finding a maximum skewness portfolio - a general solution to three-moments portfolio choice. *Journal of Economic and Dynamic Control* 28(7), 1335–1352.
- Devroye, L. and H.-K. Hwang (2006). Width and mode of the profile for some random trees of logarithmic height. *The Annals of Applied Probability* 16(2), 886–918.
- Dittmar, R. (2002). Nonlinear pricing kernels, kurtosis preference, and evidence from the cross section of equity returns. *Journal of Finance* 57(1), 369–403.
- Epps, T. W. (1979). Comovements in stock prices in the very short run. *Journal of the American Statistical Association* 74(366), 291–298.
- Fama, E. (1965). The behaviour of stock market prices. *The Journal of Business* 38(1), 34–105.
- Guidolin, M. and A. Timmermann (2008). International asset allocation under regime switching, skew and kurtosis preferences. *Review of Financial Studies* 21(2), 889–935.

- Harvey, C. and A. Siddique (2000). Conditional skewness in asset pricing tests. *Journal of Finance* 55(3), 1263–1295.
- Harvey, C. R., J. Liechty, M. W. Liechty, and P. Mueller (2010). Portfolio selection with higher moments. *Quantitative Finance* 10(5), 469–485.
- Jondeau, E. and M. Rockinger (2003). Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. *Journal of Economic Dynamics and Control* 27(10), 1699–1737.
- Jondeau, E. and M. Rockinger (2006). Optimal portfolio allocation under higher moments, european financial management. *Journal of Financial Economics* 12(1), 29–55.
- Levy, H. (1969). A utility function depending on the first three moments. *Journal of Finance* 24(4), 715–719.
- Mandelbrot, B. (1963). The variation of certain speculative prices. *The Journal of Business* 36(4), 394–419.
- Mandelbrot, B. and H. Taylor (1967). On the distribution of stock price differences. *Operations Research* 15(6), 1057–1062.
- Martellini, L. and V. . Ziemann (2010). Improved estimates of higher-order comoments and implications for portfolio selection. *The Review of Financial Studies* 23(4), 1467–1502.
- Mencía, J. and E. Sentana (2009). Multivariate location-scale mixtures of normals and mean-variance-skewness portfolio allocation. *Journal of Econometrics* 153(2), 105–121.
- Mitton, T. and K. Vorkink (2007). Equilibrium underdiversification and the preference for skewness. *The Review of Financial Studies* 20(4), 1255–1288.
- O’Hara, M. (2003). Presidential address: Liquidity and price discovery. *Journal of Finance* 58(4), 1335–1354.
- Rachev, S. and S. Mittnik (2000). *Stable Paretian Models in Finance*. Wiley, New York.
- von Mises, R. (1947). On the asymptotic differentiable statistical functions. *The Annals of Mathematical Statistics* 18(3), 105–121.
- Williams, J., J. Chen, and M. Buckle (2012). How predictable are equity covariance matrices? evidence from high frequency data for four markets. Technical report, University of Aberdeen Working Paper.

Appendix: Tables and Figures

Table 1: Comparison of realized moment estimators. The Monte-Carlo experiments are conducted for a simulated portfolio of twelve stocks. The coupling distribution uses a multivariate T distribution, with a random draw of a positive semi-definite matrix using the **randcorr** routine, one degree of freedom is assigned to each stock, therefore twelve degrees of freedom are included in the multivariate distribution. The marginal distributions are of the generalized extreme value type, with shape parameter 0.1, mean 0 and variance of unity. The estimators considered are our ‘adjusted’ needlework estimator, the needlework estimator of Audrino and Corsi (2010), a ten second fixed grid and a one minute fixed grid. The results present the average Frobenius norm of the difference between estimated comoment and the ‘true’ comoment from the underlying data generating process after 10,000 replications. The higher the number the worse the overall estimate is. The first column presents the simulation conditions. γ is the standard deviation of microstructure noise as a percentage of the variance of the marginal distribution and λ is the mean number of ticks anticipated for a days trading.

	Adjusted Needlework			Needlework			10 Second Fixed			One Minute Fixed		
	Covar	Coskew	Cokurt	Covar	Coskew	Cokurt	Covar	Coskew	Cokurt	Covar	Coskew	Cokurt
$\gamma = 0.001, \lambda = 5000$	0.11545	0.00056	0.00005	0.30713	0.11633	0.04577	0.18199	0.04934	0.01361	0.26449	0.00329	0.00028
$\gamma = 0.001, \lambda = 8000$	0.10235	0.00070	0.00005	0.23576	0.08130	0.02601	0.17087	0.04872	0.01266	0.26690	0.00158	0.00010
$\gamma = 0.005, \lambda = 8000$	0.27705	0.00105	0.00015	0.27529	0.10640	0.04049	0.20706	0.06267	0.01836	0.25351	0.00261	0.00022
$\gamma = 0.005, \lambda = 5000$	0.49048	0.00115	0.00019	0.24344	0.06232	0.02380	0.18213	0.04785	0.01303	0.27066	0.00159	0.00011
$\gamma = 0.01, \lambda = 5000$	1.29456	0.00380	0.00123	0.31299	0.12186	0.05007	0.19002	0.05510	0.01558	0.27051	0.00188	0.00013
$\gamma = 0.01, \lambda = 8000$	2.05995	0.00408	0.00168	0.34393	0.15913	0.06540	0.22678	0.08120	0.02566	0.27217	0.00685	0.00073
$\gamma = 0.1, \lambda = 8000$	128.78794	2.37114	8.23259	1.21473	0.59971	0.56745	1.33859	0.44906	0.28119	0.27492	0.00283	0.00027
$\gamma = 0.1, \lambda = 5000$	209.57251	3.24780	13.77931	0.59291	0.32621	0.16910	1.15867	0.42397	0.26270	0.26356	0.00630	0.00069

Table 2: The selected stocks in the sample. We present the company name and NYSE index. The Reuters Information Code (RIC) is used to extract the tick data from the TickHistory database. The extension (.N) reserves the ticks for only transactions on the NYSE. For our main analysis on the stability of the realized higher moments we use a sample from the start of trading January 1, 2005 to the close of trade October 10, 2011. For our example portfolio we utilize the minimum realized variance portfolio for the trading days for the two week period from September 1, 2008 to September 15, 2008. The complete dataset is available at <http://www.abdn.ac.uk/~acc138>.

Comp.	1	NYSE Code	Company	RIC Code	Available Ticks	2005-2011	Ticks September 2008	Weight in Example Portfolio
Comp.	2	AA	Alcoa Inc	AA.N	12,950,641	195,177	-0.031	
Comp.	3	AXP	American Express Co	AXP.N	14,822,386	243,517	-0.050	
Comp.	4	BA	Boeing Co	BA.N	12,022,000	194,491	-0.018	
Comp.	5	CAT	Caterpillar Inc	CAT.N	13,512,817	185,598	0.011	
Comp.	6	CVX	Chevron Corp	CVX.N	19,785,954	284,163	0.126	
Comp.	7	DD	Du Pont De Nemours And Co	DD.N	10,743,637	143,733	0.020	
Comp.	8	XOM	Exxon Mobil Corp	XOM.N	12,945,826	185,182	-0.060	
Comp.	9	GE	General Electric Company	GE.N	22,340,767	555,128	0.200	
Comp.	10	HPQ	Hewlett Packard Co	HD.N	15,070,811	266,129	-0.012	
Comp.	11	HD	Home Depot Inc	HPQ.N	16,721,162	246,379	0.074	
Comp.	12	IBM	International Business Machines Co...	IBM.N	15,475,512	229,206	0.029	
Comp.	13	JNJ	Johnson & Johnson	JNJ.N	16,008,125	234,545	0.215	
Comp.	14	KO	Coca Cola Co	KO.N	13,663,117	212,828	0.149	
Comp.	15	MCD	McDonalds Corp	MCD.N	12,348,177	201,398	0.068	
Comp.	16	PFE	Pfizer Inc	PFE.N	16,257,888	258,440	0.045	
Comp.	17	PG	Procter & Gamble Co	PG.N	15,441,131	243,177	0.188	
Comp.	18	TRV	Travelers Companies Inc	T.N	16,213,077	273,027	-0.037	
Comp.	19	UTX	United Technologies Corp	UTX.N	10,796,220	172,741	-0.034	
Comp.	20	VZ	Verizon Communications Inc	VZ.N	13,944,319	246,131	0.008	

Table 3: Distribution of standardized tick times and returns for the complete cross section of the sample of 20 stocks chosen for the empirical study.

	Mean ticks	Median ticks	Std ticks	Min ticks	Max ticks	Mean zeros	Median zeros	Std zeros	Min zeros	Max zeros
2005	4340.6	4316.0	527.9	2151.0	5860.0	7779.4	7800.5	1177.1	3868.0	10913.0
2006	5522.4	5357.0	897.4	2662.0	8547.0	10951.1	10150.0	2674.7	5563.0	22568.0
2007	10862.0	10252.0	2859.7	3721.0	24422.0	21087.2	20726.0	6109.2	6584.0	48318.0
2008	12717.9	11838.0	4226.5	4032.0	31546.0	15118.6	14543.0	3838.9	6366.0	26718.0
2009	10514.3	9680.0	3013.3	2705.0	19597.0	22941.6	22345.0	5352.2	7386.0	42747.0
2010	7787.1	6983.0	2806.4	2663.0	24797.0	19438.7	18844.5	5065.4	5849.0	40979.0
2011	7563.8	6460.0	3311.5	4065.0	25805.0	17440.8	16604.0	4657.4	8254.0	37696.0

Table 4: For the raw moments this is the sum through time of the super diagonal elements of the moment-comoment arrays $\mathbf{M}_{\{1,\dots,r\}}$ for each year. I.e. the cumulative variation for the first four moments for the cross section of stocks. The total variation is resampled, with replacement and a minimum, maximum and mean are computed.

	Max \mathbf{m}_2	Mean \mathbf{m}_2	Min \mathbf{m}_2	Max \mathbf{m}_3	Mean \mathbf{m}_3	Min \mathbf{m}_3	Max \mathbf{m}_4	Mean \mathbf{m}_4	Min \mathbf{m}_4
2005	0.26765	0.19531	0.11465	2.92683	0.05708	-4.73045	4.45881	2.59540	1.60031
2006	0.23866	0.17428	0.12426	3.52920	0.18451	-2.27888	3.63199	2.49740	1.87339
2007	0.25704	0.14609	0.10616	2.21323	0.09786	-2.60722	6.18327	2.59650	1.64472
2008	0.77689	0.25469	0.15161	8.50700	-0.04109	-8.27013	18.99525	4.54911	2.38020
2009	0.35590	0.22297	0.13192	5.21721	0.16490	-4.85769	8.09817	3.76480	1.68060
2010	0.38690	0.16912	0.12714	2.88513	0.22058	-2.71897	11.00969	2.55732	1.61543
2011	0.29538	0.17439	0.12631	3.84122	0.15666	-3.50873	6.04316	2.61125	1.79309

Table 5: Annualized and normalized off-diagonal averages for the cross section of selected 20 stocks from the Dow 30 components. This is computed for the $\mathbf{R}_{\{1,\dots,r\}}$ arrays, using the annualized averages from Table 4. The *offdiag* operator stacks the off diagonal elements i.e $\mathbf{r}_r = \text{offdiag}\mathcal{R}_r$, where $\mathbf{r}_r = [r_{i,j,k,l,\dots,[r]}]_{i \neq j \neq k \neq l \dots, [r]}$.

	Max \mathbf{r}_2	Mean \mathbf{r}_2	Min \mathbf{r}_2	Max \mathbf{r}_3	Mean \mathbf{r}_3	Min \mathbf{r}_3	Max \mathbf{r}_4	Mean \mathbf{r}_4	Min \mathbf{r}_4
2005	0.42414	0.25623	0.00000	0.65583	0.07122	-0.01946	0.26551	0.05535	0.00000
2006	0.42326	0.22958	0.09395	1.17802	0.07265	-0.14208	0.35044	0.04455	0.00432
2007	0.45841	0.26167	0.10391	0.42513	0.04150	-0.02579	0.18766	0.04013	0.00455
2008	0.61714	0.36540	0.20988	1.40010	0.08522	-0.04821	0.34985	0.08939	0.01562
2009	0.52202	0.35724	0.17779	0.55300	0.08281	-0.00992	0.44887	0.09636	0.02665
2010	0.63138	0.36187	0.11645	0.52925	0.08161	-0.01595	0.40972	0.10685	0.01067
2011	0.67797	0.36436	0.13265	0.95963	0.10449	-0.06784	0.45625	0.11590	0.00920

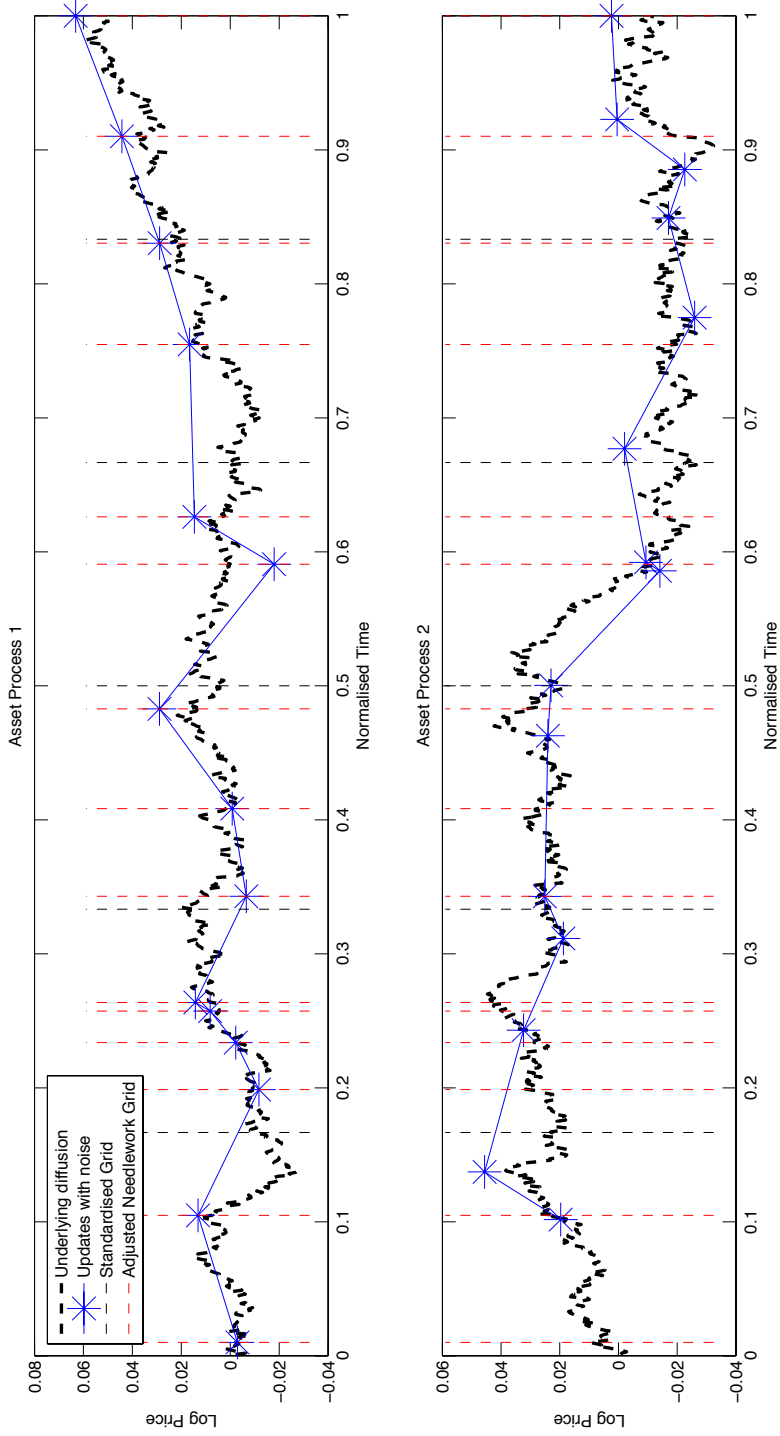


Figure 1: Comparison of our adjusted needlework estimator versus a fixed grid. The black dashed lines represent the underlying diffusion with realized increments δ_t . The market updates at time points marked by the blue markers. To compare the two processes we need to interpolate each price process to a master grid. The black lines illustrate a uniform grid, which is the same for each stock. The red lines represent the adjusted needlework estimators sampling times, based on the least updated stock over a short fixed interval. For the set of stocks this should capture, cross sectionally, the largest number of price movements. However the effectiveness of the sampling scheme is reduced by large amounts of microstructure noise and very low frequency updating, see Table 1.

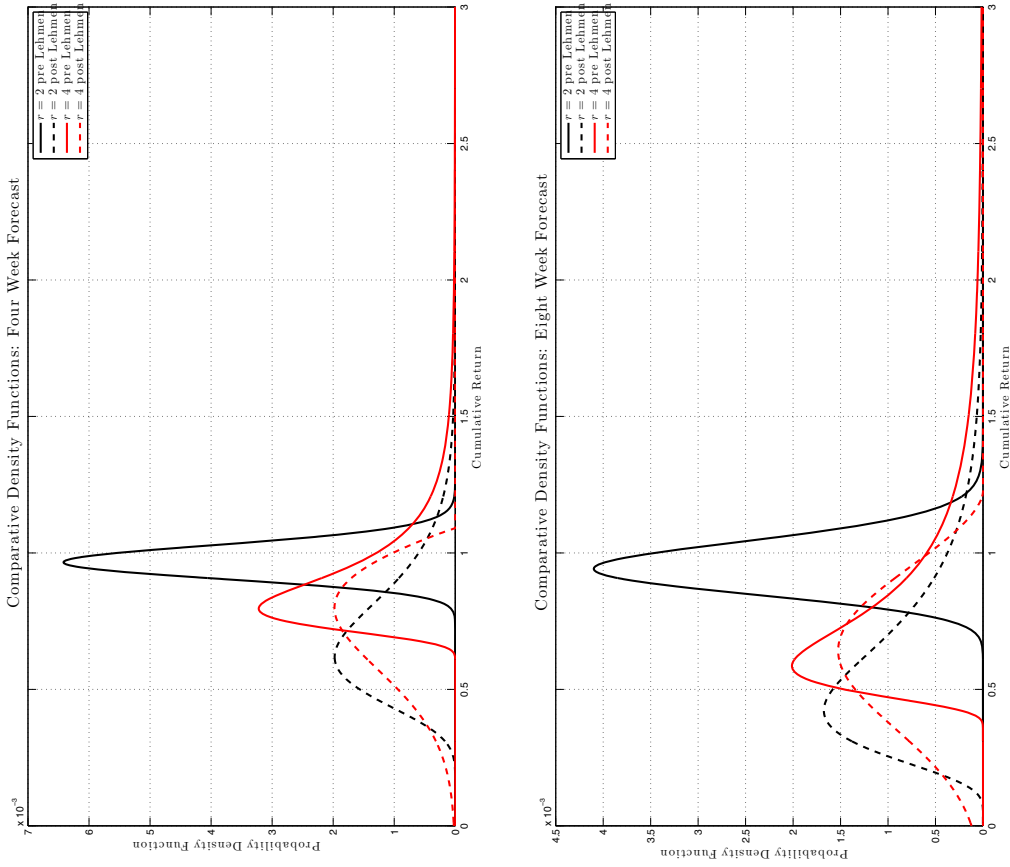


Figure 2: Comparison of two versus four moment density functions.

Tick by tick data for twelve stocks pre and post September 16, 2008 is used to impute the forecasted versus realized distribution measure by a two and four moment portfolio density approach. The portfolio is computed from the minimum variance weights (with short-selling) using the formula $\omega = \mathbf{e}\mathbf{M}_2^{-1}(\mathbf{e}'\mathbf{M}_2^{-1}\mathbf{e})^{-1}$, where \mathbf{M}_2 is the estimated variance covariance matrix for the preceding quarter. The top plot is for a four week window, pre and post September 16, 2008 and the bottom plot is for an eight week window pre and post September 16, 2008. For the two moment system returns are assumed to be from a Gumbel distribution therefore portfolio densities are from a Weibull distribution instead of log-normal. Several other distributions are available that can be fitted by two moments with a deformation parameter, for instance the Skew normal. With four moments we can expand the class of candidate distributions to include more generalized extreme value and other types of Stable Paretian distributions, in this case a generalized extreme value distribution is used for the fitted characteristic function from the raw moments.

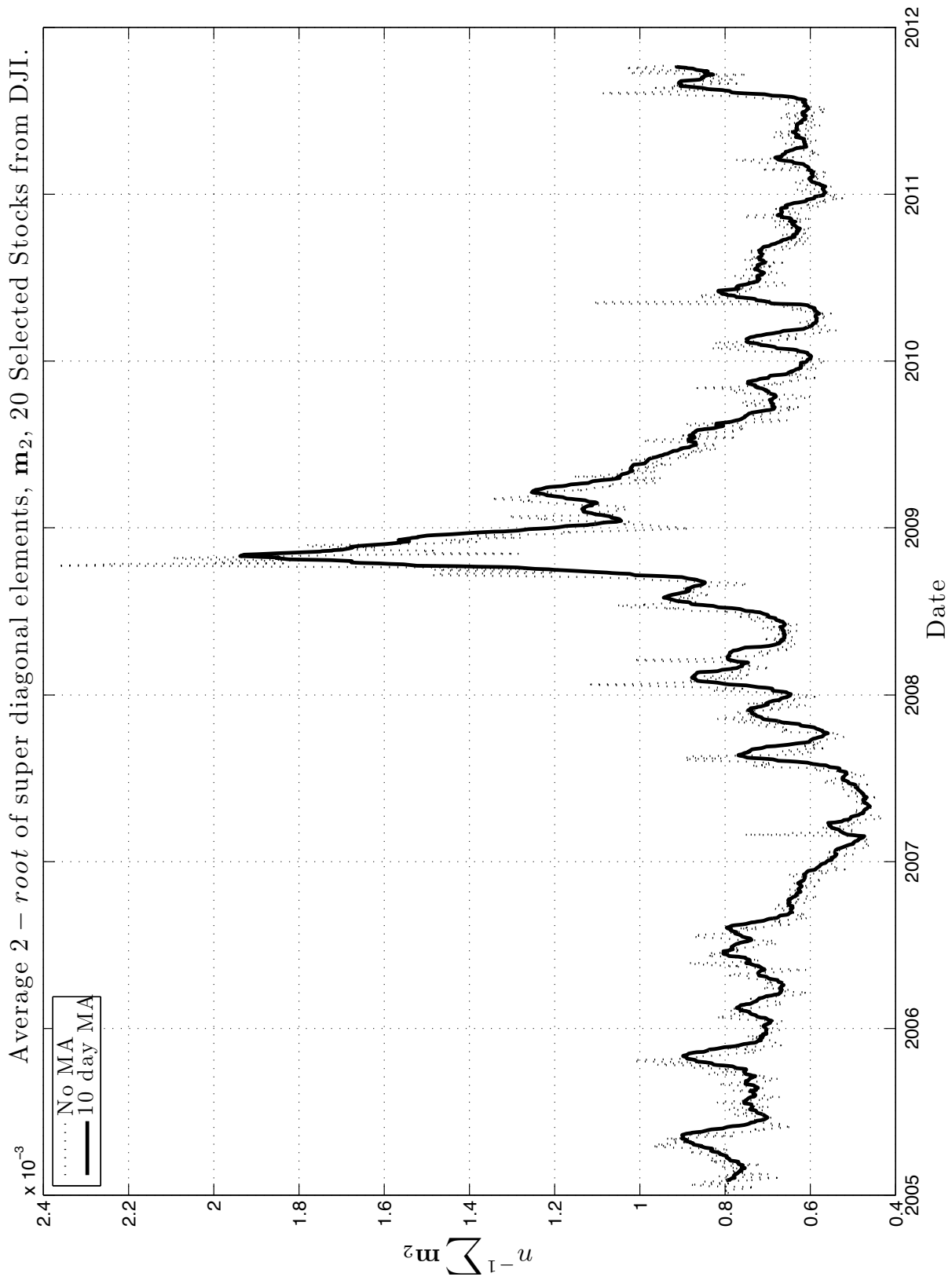


Figure 3: Cumulative daily $2 - \text{root}$ of the diagonal realized variances for 20 stocks from the Dow 30 Index components.

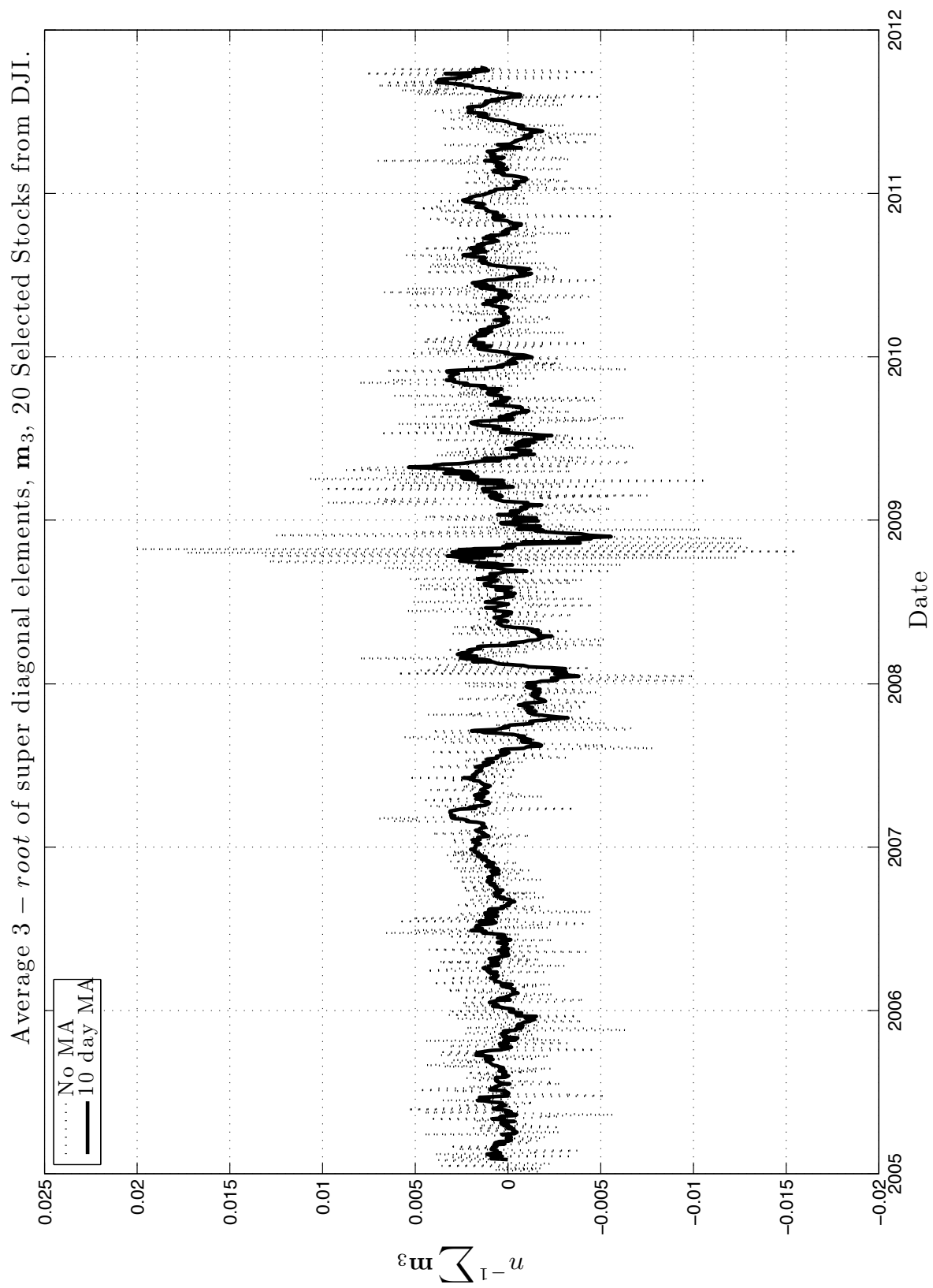


Figure 4: Cumulative daily 3 – root of the diagonal realized raw skewness for 20 stocks from the Dow 30 Index components.

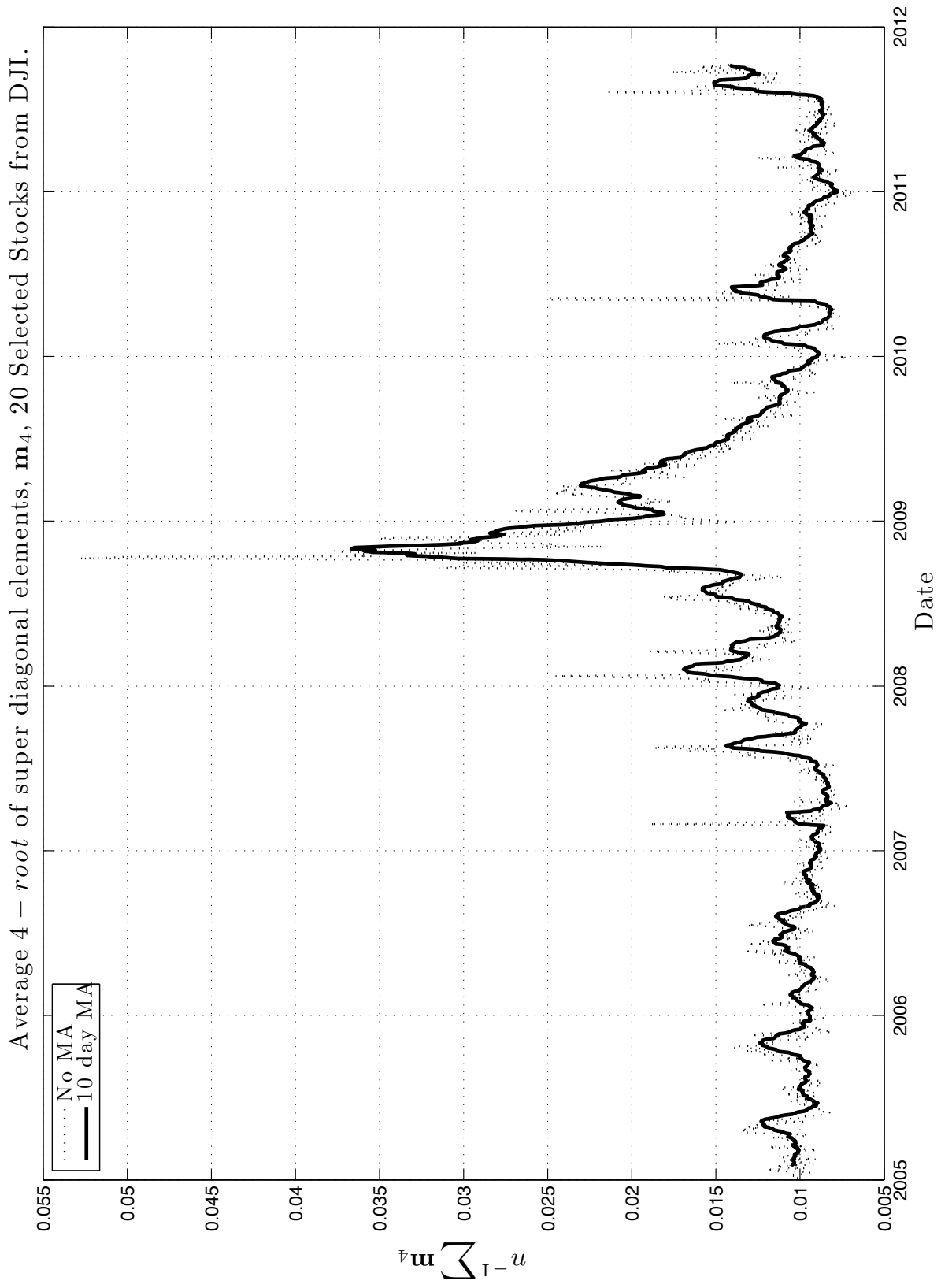


Figure 5: Cumulative daily $4 - \text{root}$ of the diagonal realized raw kurtosis for 20 stocks from the Dow 30 Index components.

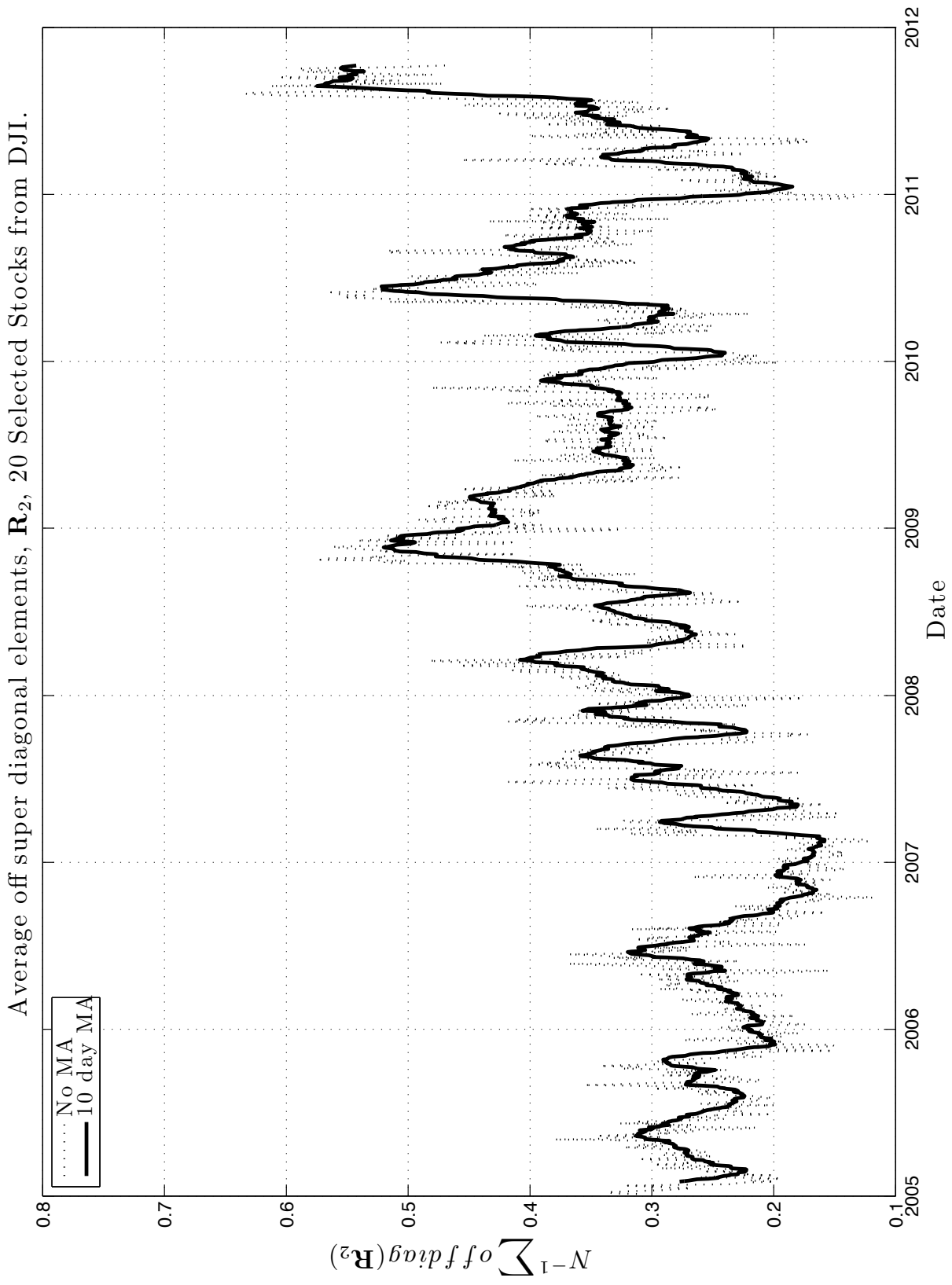


Figure 6: Average of daily cumulative normalized off-diagonal elements $\text{offdiag} \mathbf{R}_2$ for 20 selected stocks from the Dow 30 Index components. N is the number of off super diagonal elements, for $n^2 - n$.

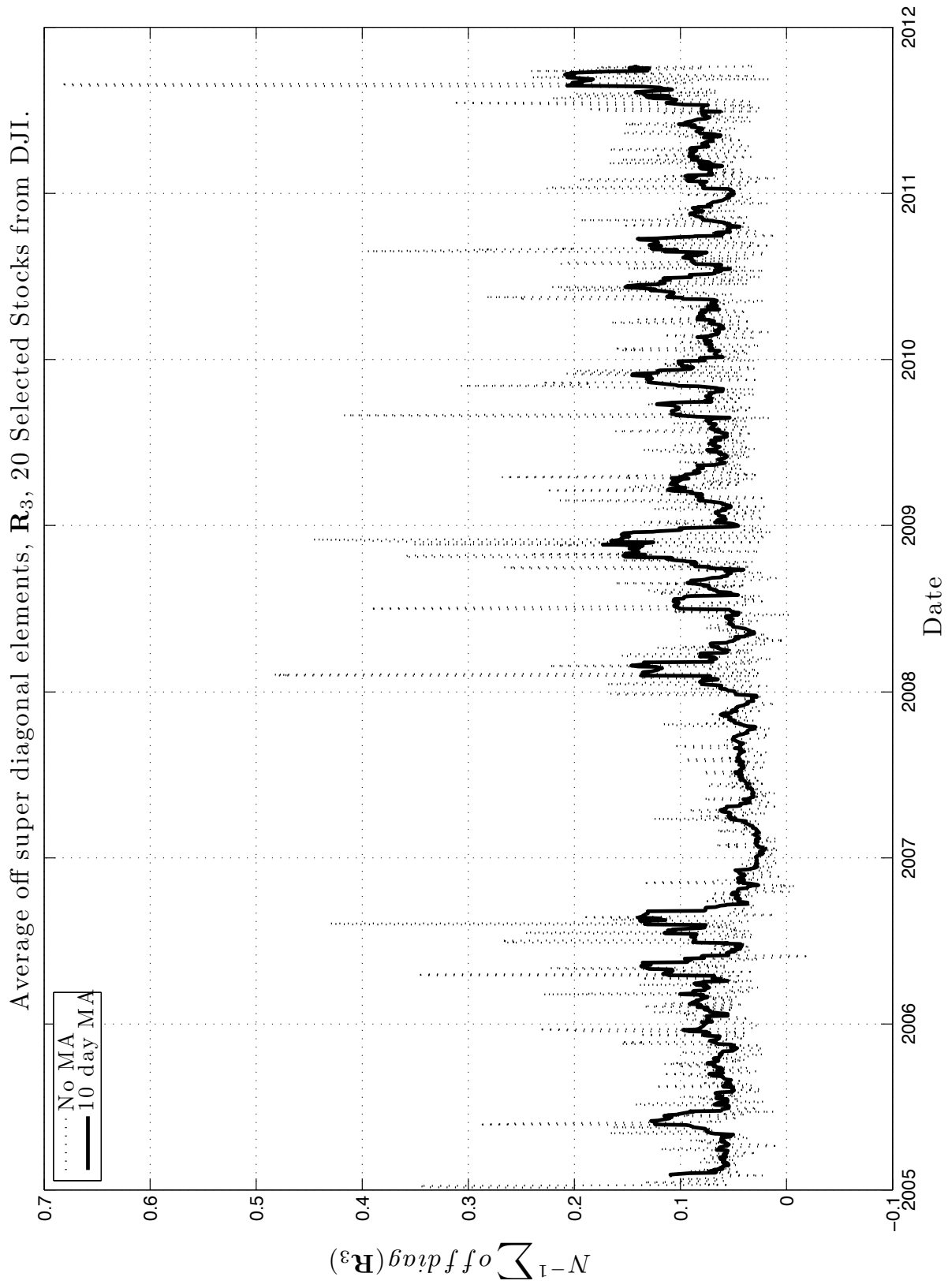


Figure 7: Average of daily cumulative normalized off-diagonal elements of $fdiag\mathbf{R}_3$ for 20 selected stocks from the Dow 30 Index components. N is the number of off super diagonal elements, for $n^3 - n$.

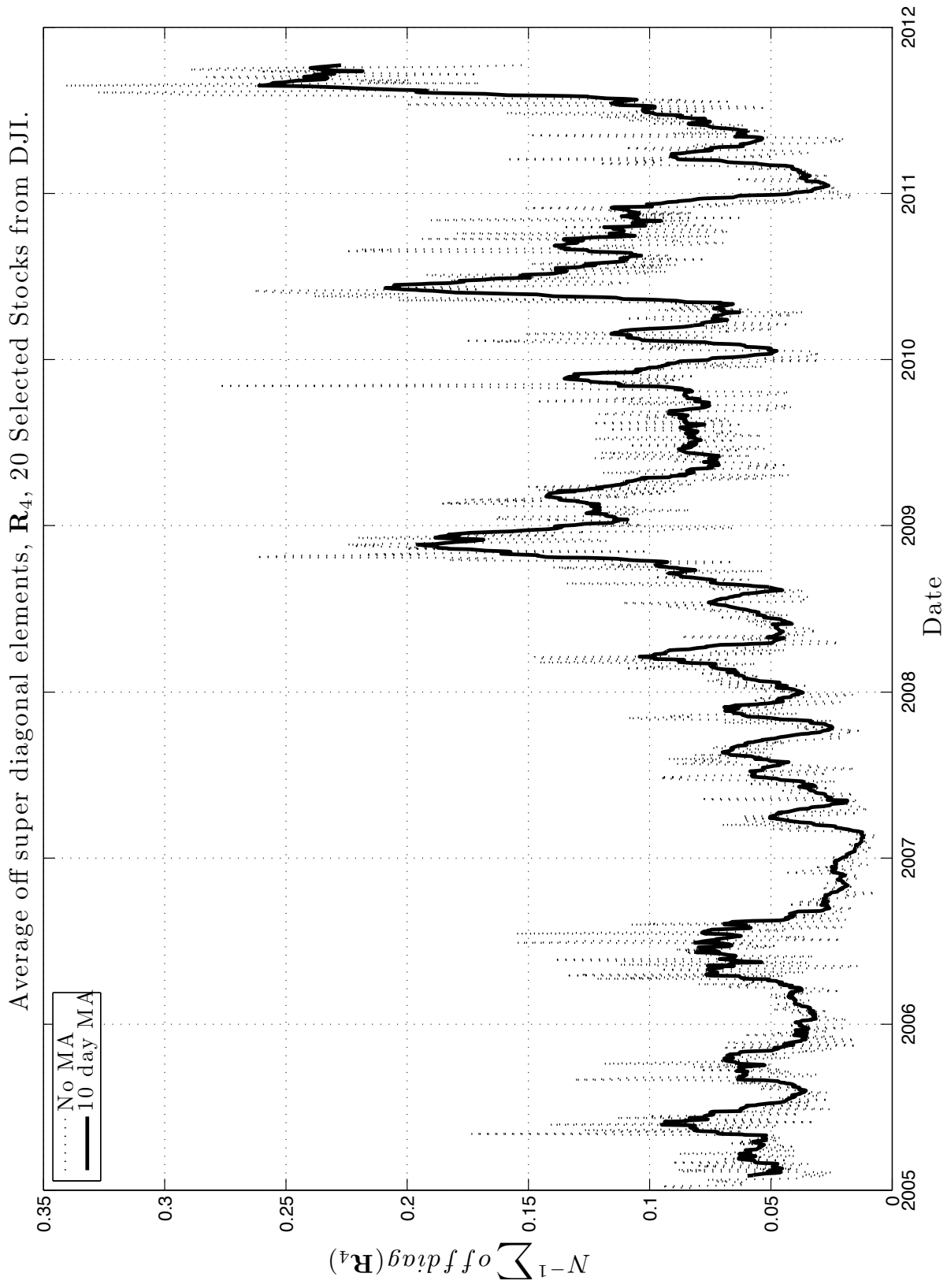


Figure 8: Average of daily cumulative normalized off-diagonal elements of $fdiag\mathbf{R}_4$ for 20 selected stocks from the Dow 30 Index components. N is the number of off super diagonal elements, for $n^4 - n$.